Calculating Physicochemical Properties for Environmental Modeling Using SPARC

L. A. Carreira[a]

T.S. Whiteside[a],S.N. Ayyampalayam[a] and S.H. Hilal[b]

[a]Department of Chemistry
University of Georgia

[b]U.S. Environmental Protection Agency

SPARC

SPARC

Performs

Automated

Reasoning in

Chemistry

Computer program to estimate physical properties and chemical reactivity parameters of organic compounds strictly from molecular structure

Two Modeling Approaches

Reaction at a center

- Ionization pKa
- Hydration Constant
- Tautomeric Equilibria
- Hydrolysis Rate

- The first two model the ΔG difference of the products and reactants
- The third is a sum of site reactions
- The last is the ΔG difference of the reactants and a transition state

Whole molecule

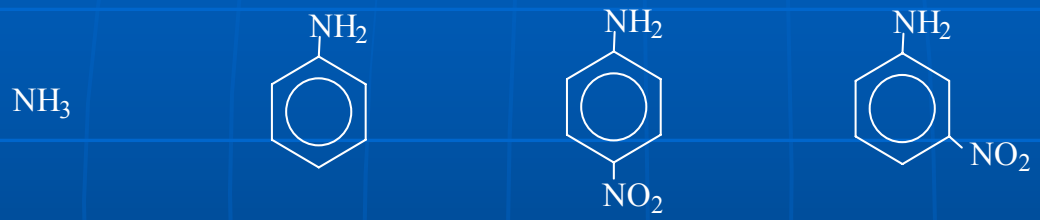
- Vapor pressure
- Boiling point
- Volume / Density
- Electron Affinity
- Polarizability / Ref. Ind.
- Henry's constant
- Activity Coefficient
- Solubility
- Distribution Coefficient

Reaction at a center

- Inherent reactivity of the stripped center
  e.g. pKa of NH_3 for ionization pKa of amines
- Resonance (PMO charge distribution)
- Sigma induction (bond electronegativity difference)
- Field effects (both direct and indirect)
- Differential solvation (steric access to charge site)
- Intramolecular hydrogen bonding

Resonance Effect

Resonance involves the delocalization of π electrons into or out of C



C replaced by a surrogate electron donor describing delocalization of NBMO electrons into or out of C_i by the R_π network

$$\delta\,(\Delta E)_{res} = \rho_{res}\,\Delta q_c$$

Field Effect Model

Charges or electric dipoles in S interacting with charges or dipoles in C

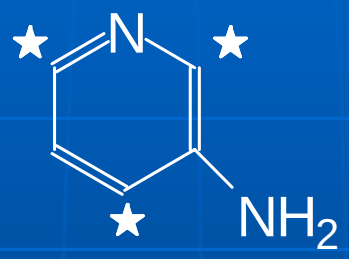
$$\delta\ (\Delta E)_{field} = \frac{\delta q_c q_s}{r'_{cs} D_e} + \frac{\delta q_c \mu_s Cos\ \theta_{cs}}{r^2_{cs} D_e}$$

$$+ \frac{\delta D_c q_S}{r^2_{cs} D_e} + ...$$

$$\delta\ \Delta E_{field} = \rho_{ele} \sigma_{cs} F_s$$

Mesomeric Field Effect

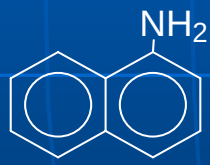
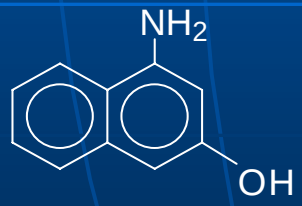
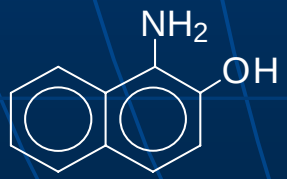
S induced electrical fields in R_π which interact with the charges in C



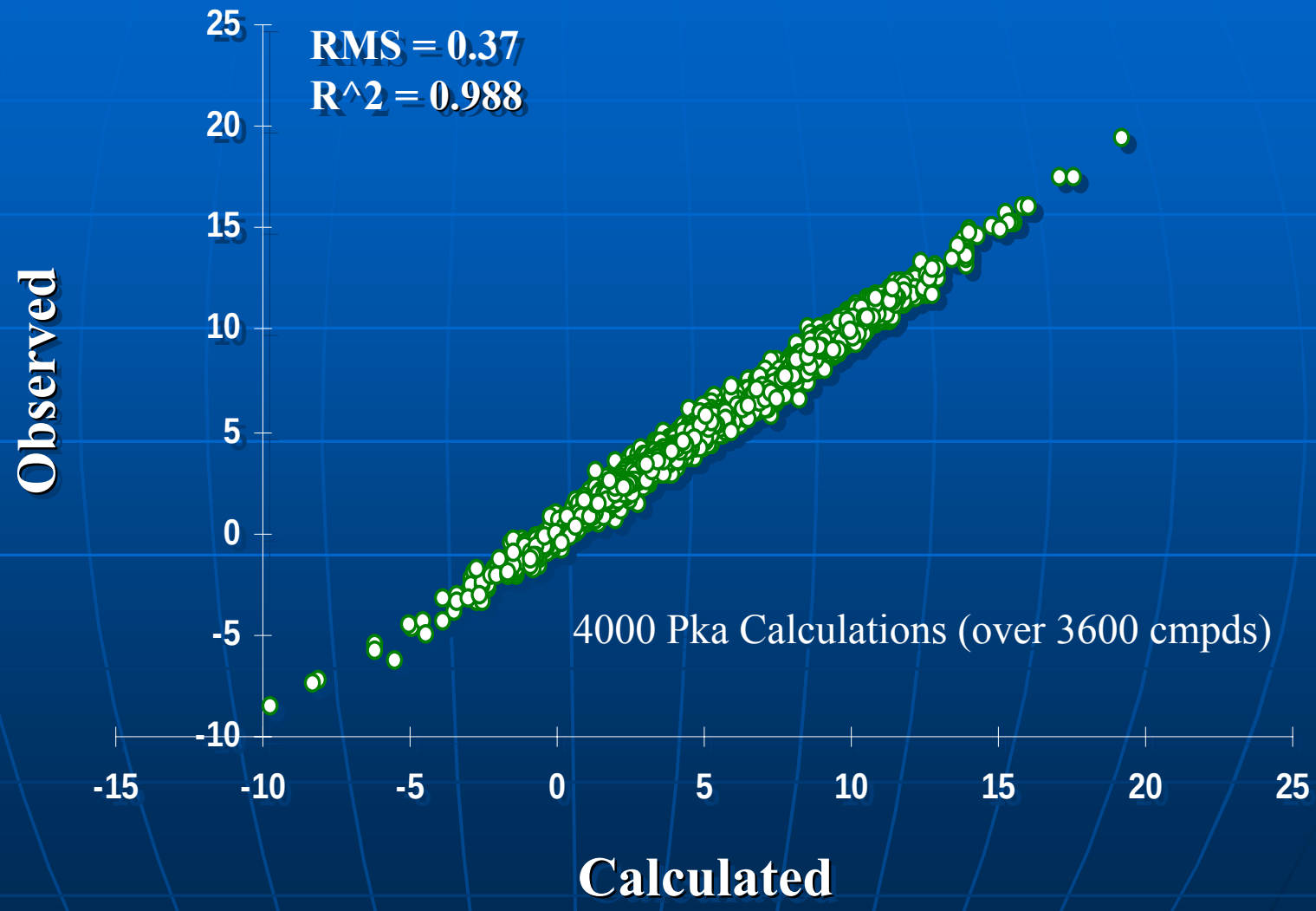
A collection of discrete charges, $q_{R,}$ with the contribution of each described by the Field mutipole expansion

$$\delta\ \Delta E_{MF} = \rho_{ele} \sum_{k} \frac{q_{ik}}{r_{kc}} MF_s$$

Perturber Effects on pK_a

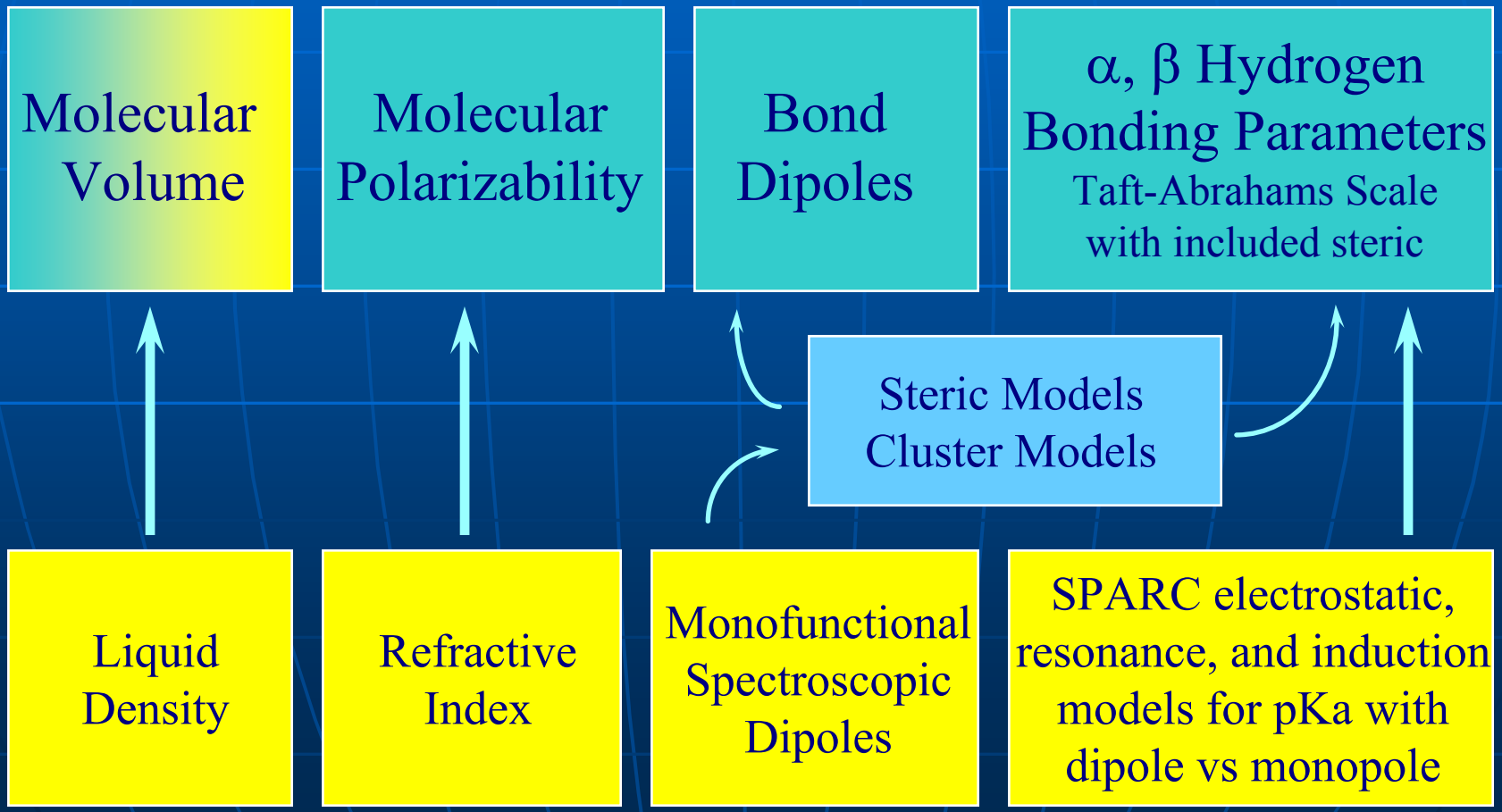
NH_3	9.8	Unperturbed
CH_3NH_2	10.5	Sigma Induction
	4.1	Resonance
	3.9	+ Steric and Solvation
	3.3	+ Electrostatic
	4.5	H-Bonding

SPARC pKa Performance

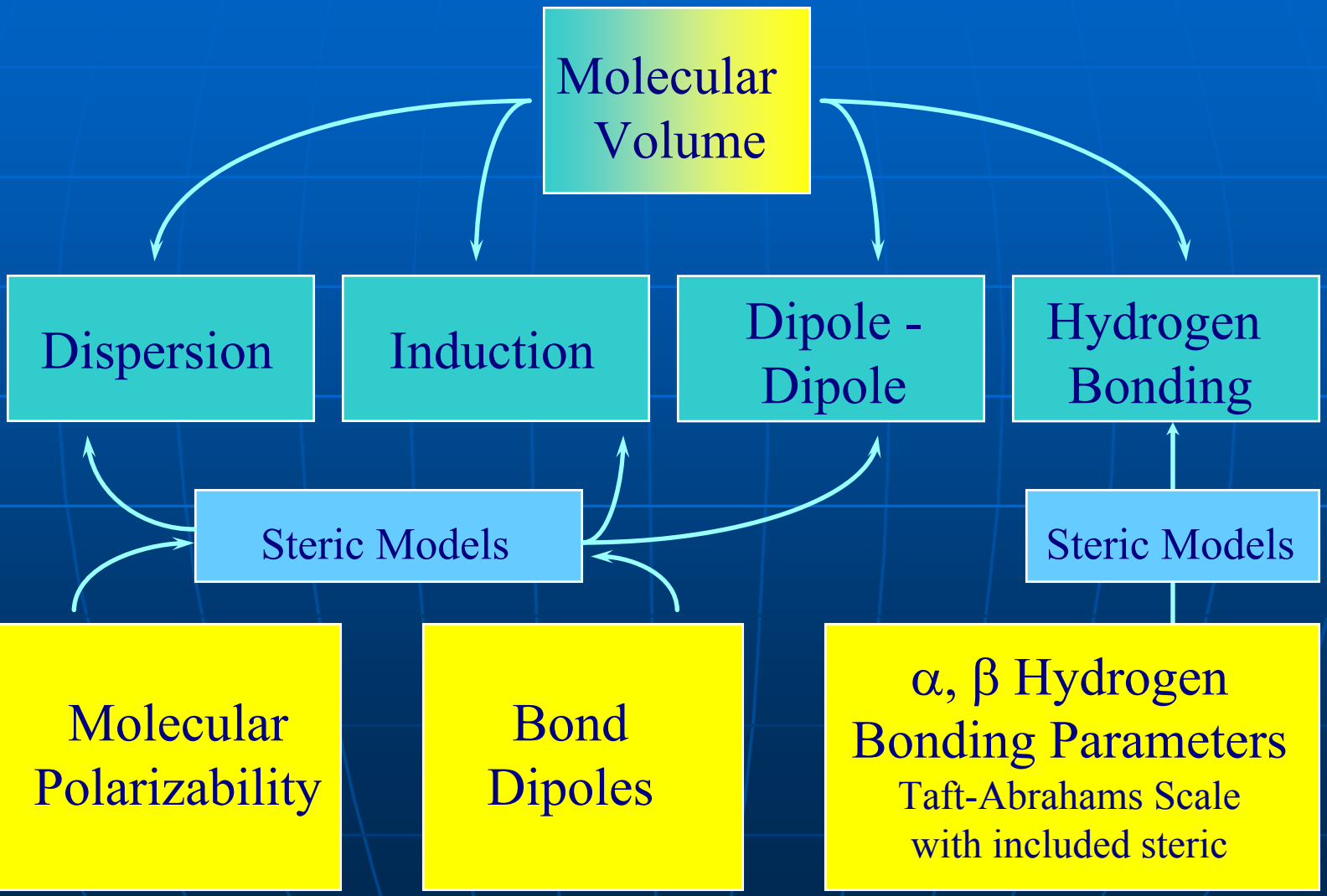


Modeling Approach for Whole Molecule Physical Properties

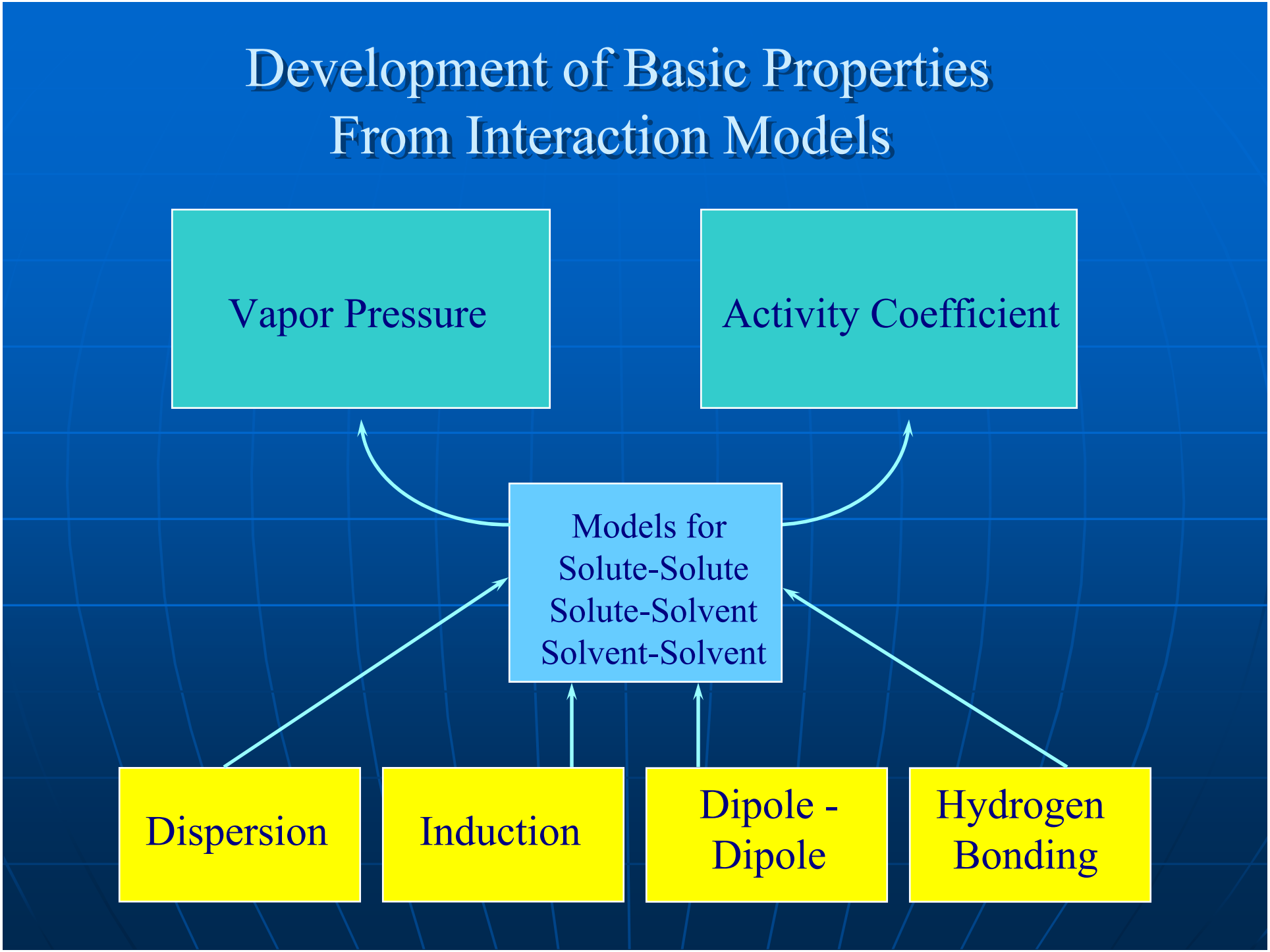
Development of Molecular Descriptors That Are Tied to Molecular Observables



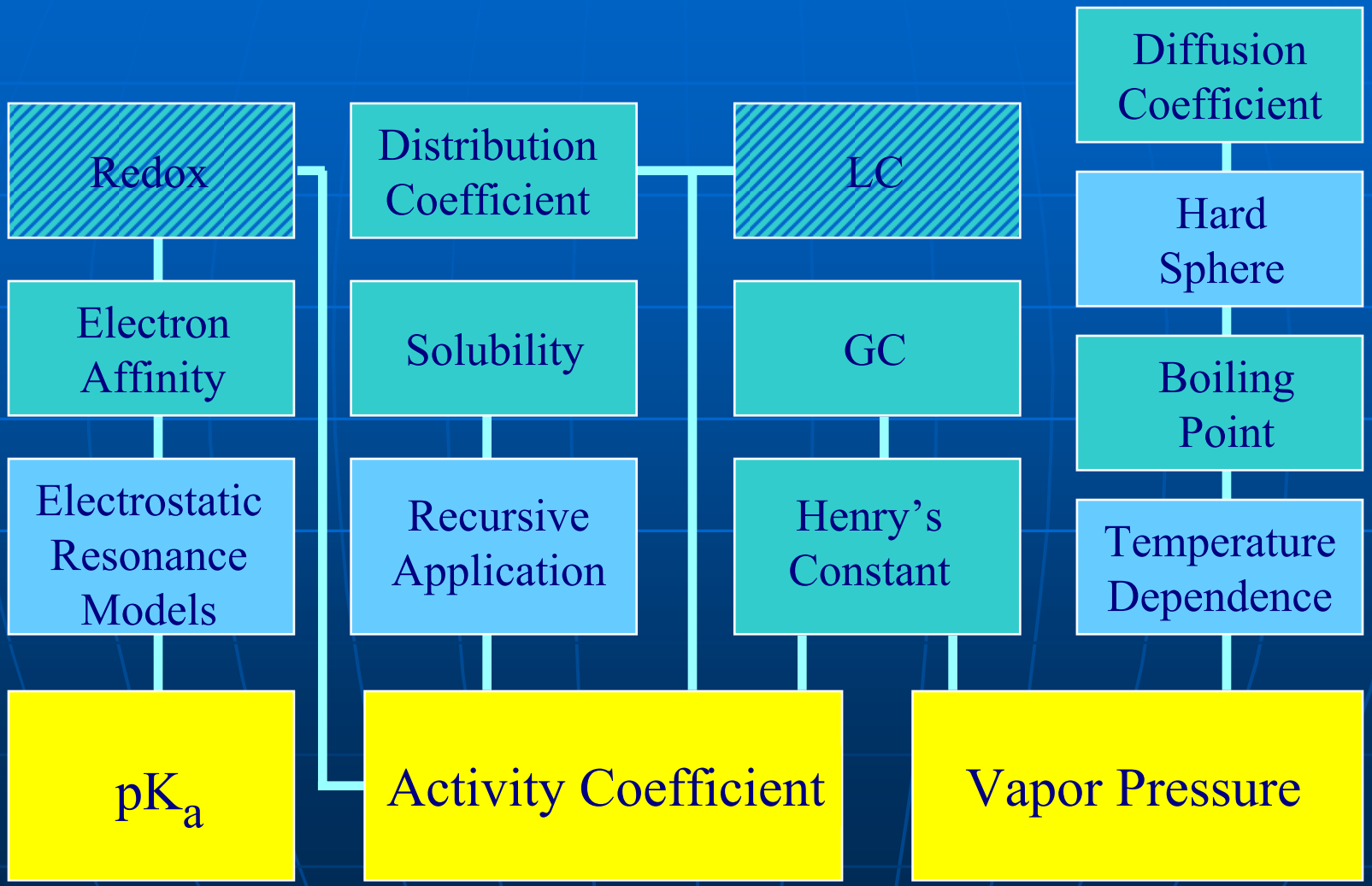
Development of Interaction Models From Molecular Descriptors



Development of Basic Properties From Interaction Models



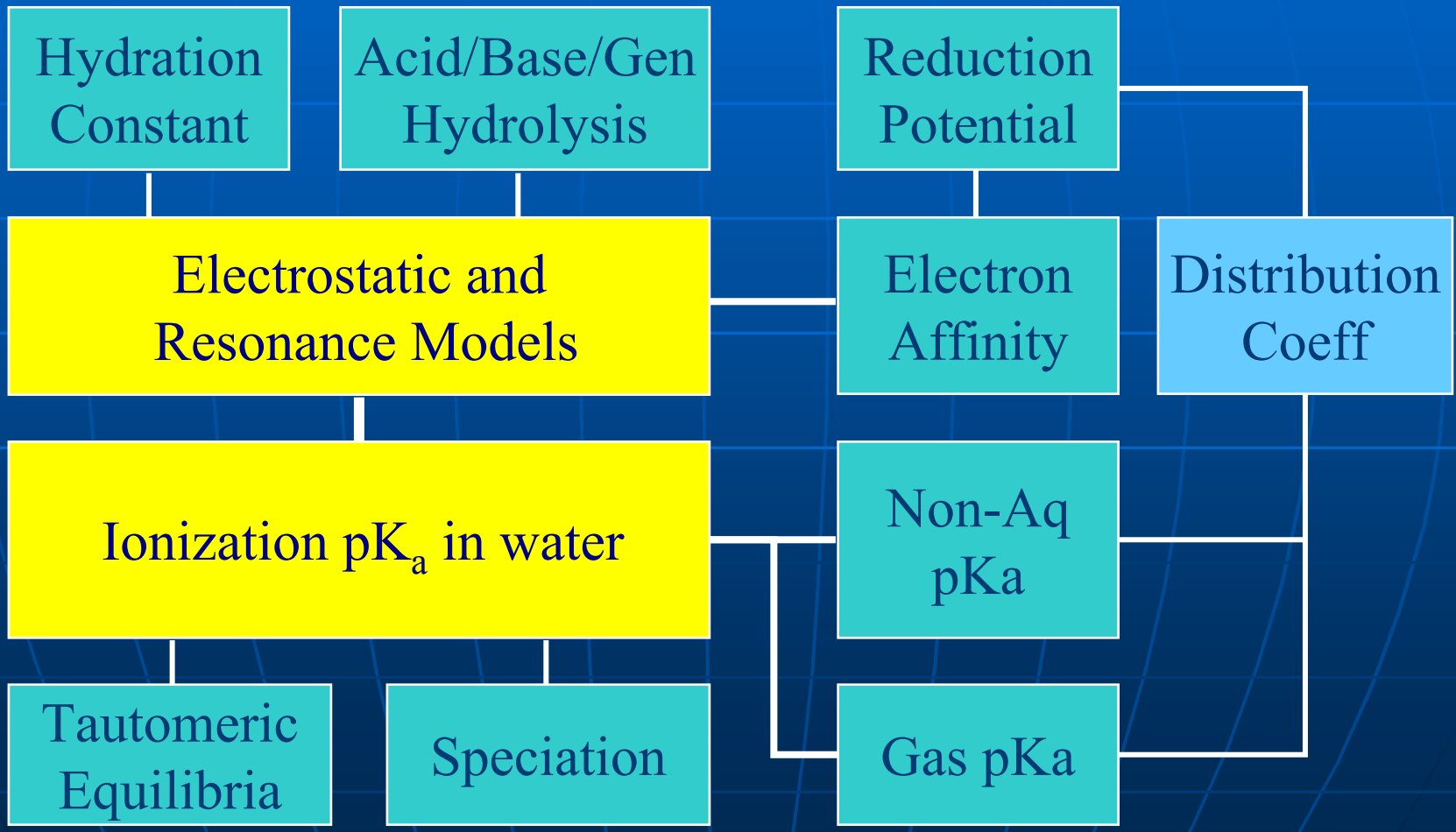
Development of General Physical Properties From the Basic Properties



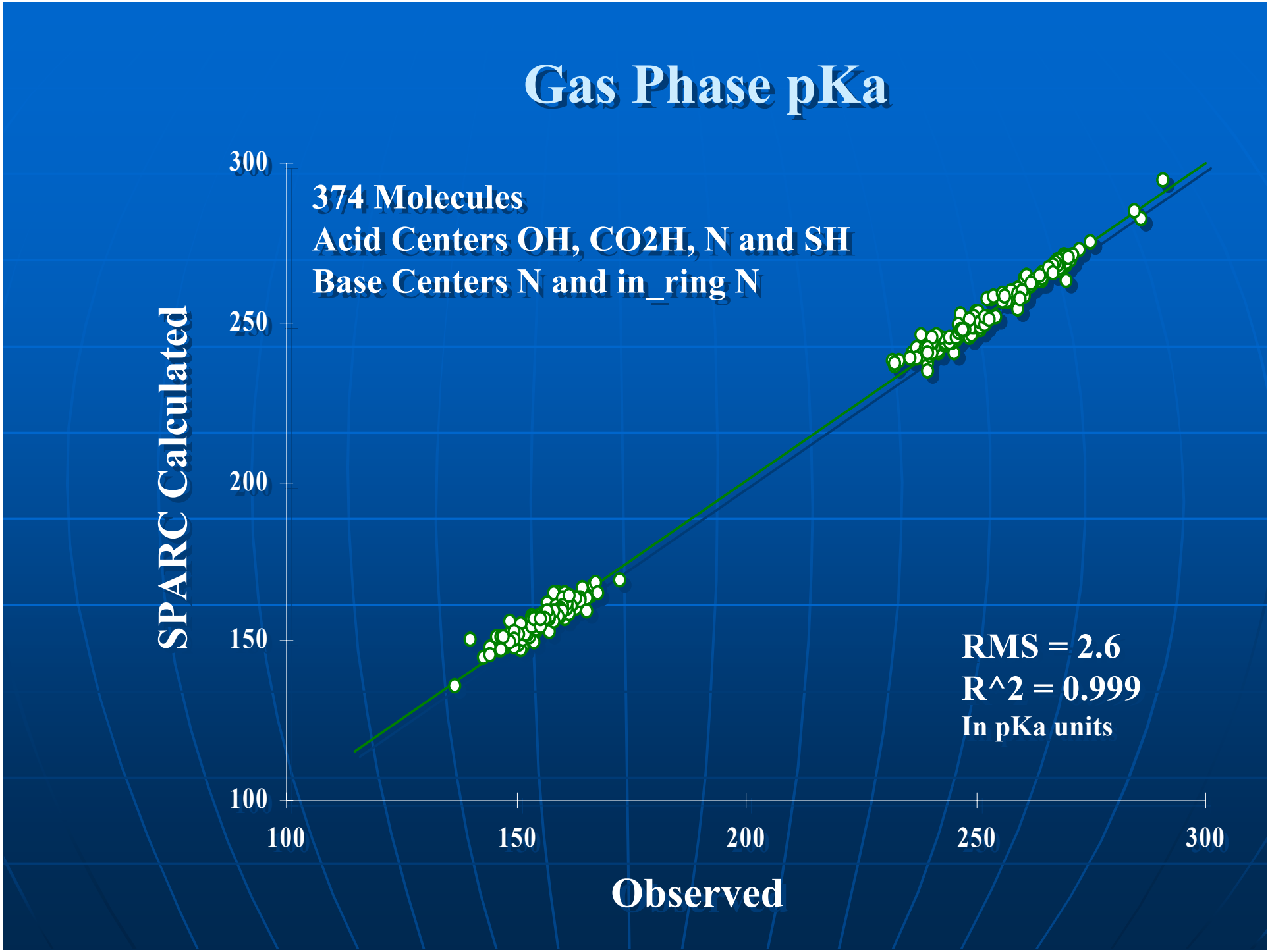
Property	Units	# Molec	RMS	Temp/Solv
Refractive Index	na	578	0.007	25
Volume	cm^3/m	1440	1.96	25-BP
Vapor Pressure	Log atm	747	0.15	25
Boiling Point	C	4000	5.7	0.1-1520 T
Heat of Vap.	Kcal/m	1263	0.30	25, BP
Diffusion Coeff.	cm^2/sec	108	0.003	25
Activity Coeff.	Raolt	2645	0.27	25-163 solv
Solubility	log MF	707	0.49	25-21 solv
Distribution Coeff	unitless	698	0.43	25-6 solv
Henry's Const.	unitless	811 360	0.41 0.15	water hexadecane

Property	Units	# Molec	RMS	Temp/Solv
GC Retention	Kovats	295	10.3	25-190 3S
LC Retention	Relative	125	0.01	25 2S
Gas pKa Non Aq pKa water pKa	Kcal Kcal pKa	400 455 4677	3.6 1.3 0.46	25 25 9S 25-100
Electron Affinity	e.v.	260	0.15	Gas 25
Carboxylic ester Phosphate ester	Log rate	1470 397	0.37 0.46	25-130 6S 25-180 W
Tautomeric Const	Log K	36	0.3	25 W
Hydration const	Log K	27	0.4	25 6S
$E_{1/2}$ reduction	e.v.	352	0.18	25-21 solv
Heat of formation	Kcal	793	1.2	25 gas/liq

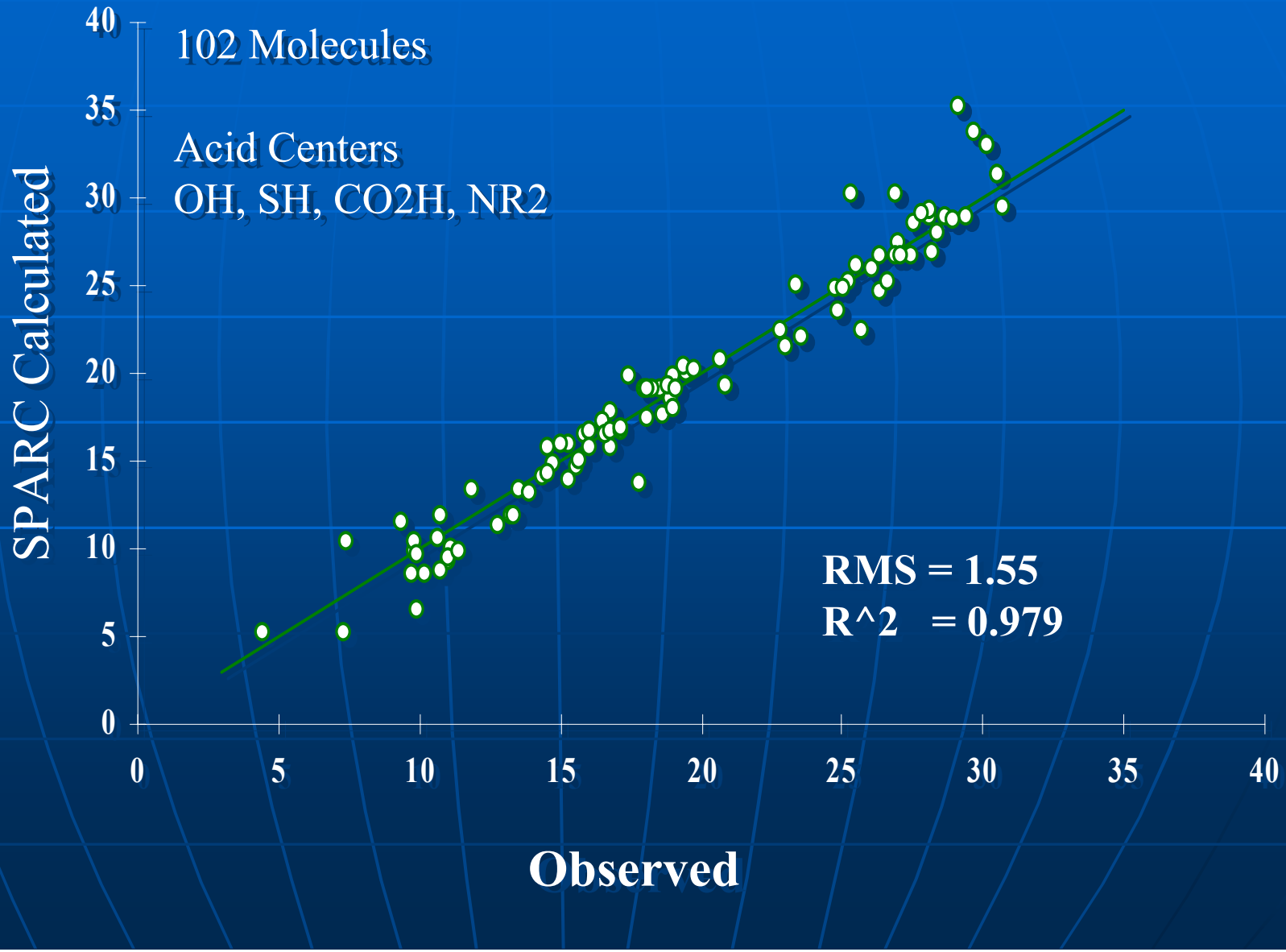
Development of General Reaction Constants From Ionization pKa and Basic Properties



Gas Phase pKa



pKa in DMSO


102 Molecules
Acid Centers
OH, SH, CO2H, NR2
SPARC Calculated
Observed
0
5
10
15
20
25
30
35
40
RMS = 1.55
R^2 = 0.979

Automate Speciation

Design the SPARC system to aid the user in *asking the right question.*

Complex Speciation in pKa

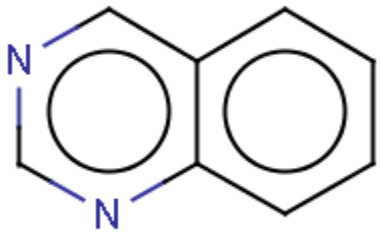
Tautomeric conversion

Hydration Reaction

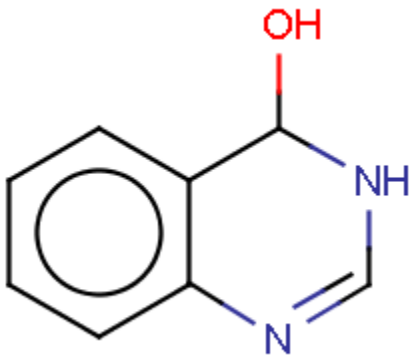
Hydrolysis Reaction

Automate SPARC so as to include the effects of several mechanisms when calculating speciation. An example where ionization, hydration and tautomeric equilibrium all play important roles will be presented.

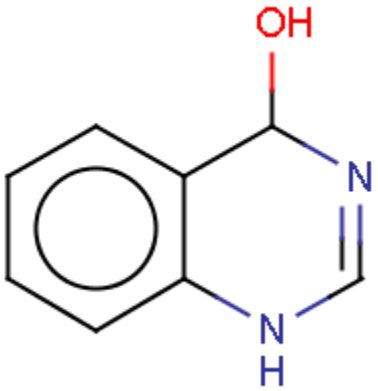
Speciation of Quinazoline



Hydrated products of Quinazoline

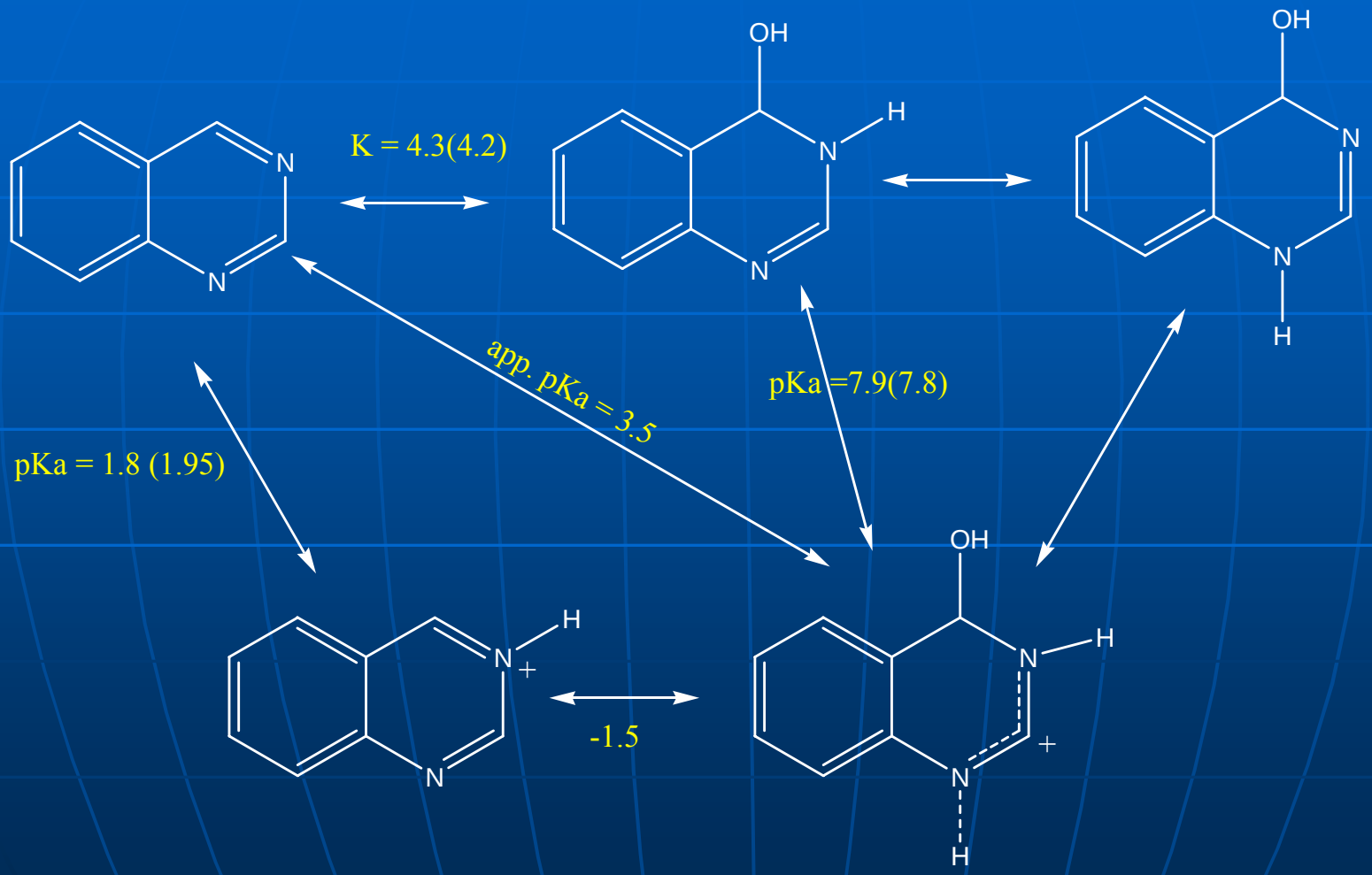


pK_{hyd} = 4.24
(4.25)
Most stable

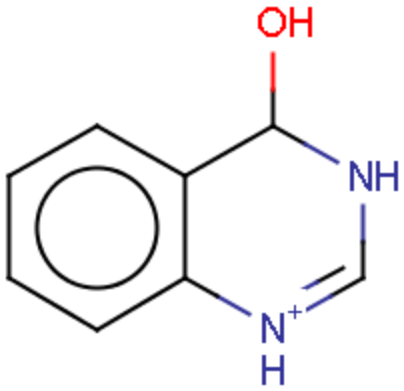


pK_{hyd} = 6.28

Hydration / Ionization

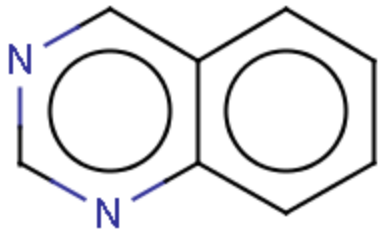


Macro pKa of Quinazoline

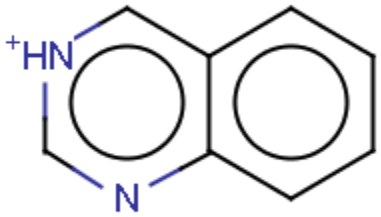


Max Fraction = 0.99

Macro Pka = 3.70

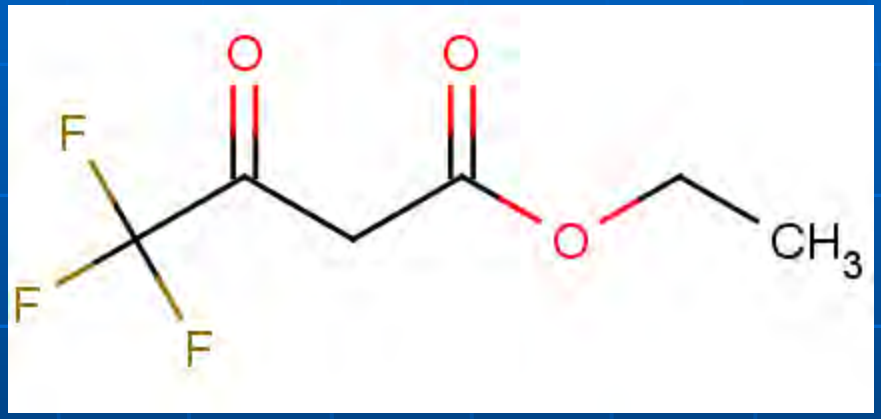


Max Fraction = 1.00



Max Fraction = 0.01

Butanoic acid, 4,4,4-trifluoro-3-oxo-, ethyl ester



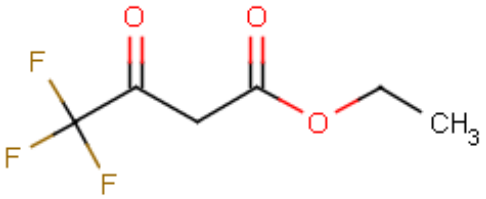
Observed at pH = 5-6

Ketone form = 7.77%

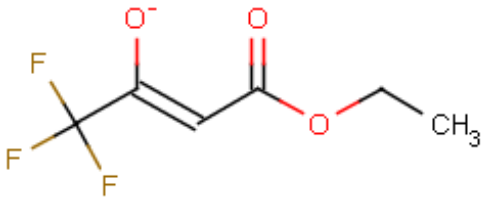
Enol form = 34.97%

Hydrate = 57.26%

Macro pKa

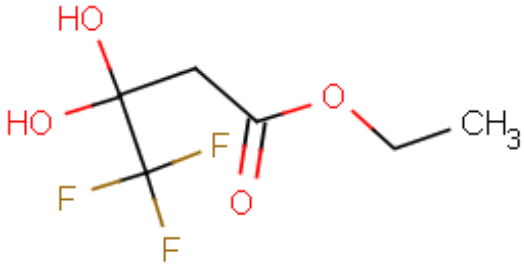


Max Fraction = 0.07

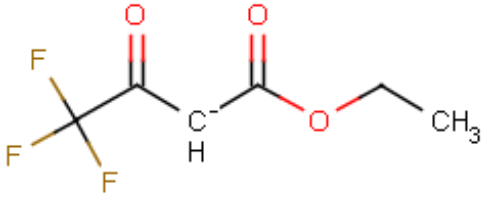


Max Fraction = 0.33

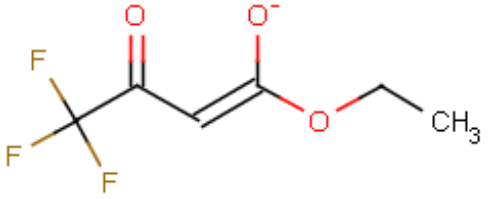
Macro Pka = 5.68



Max Fraction = 0.92

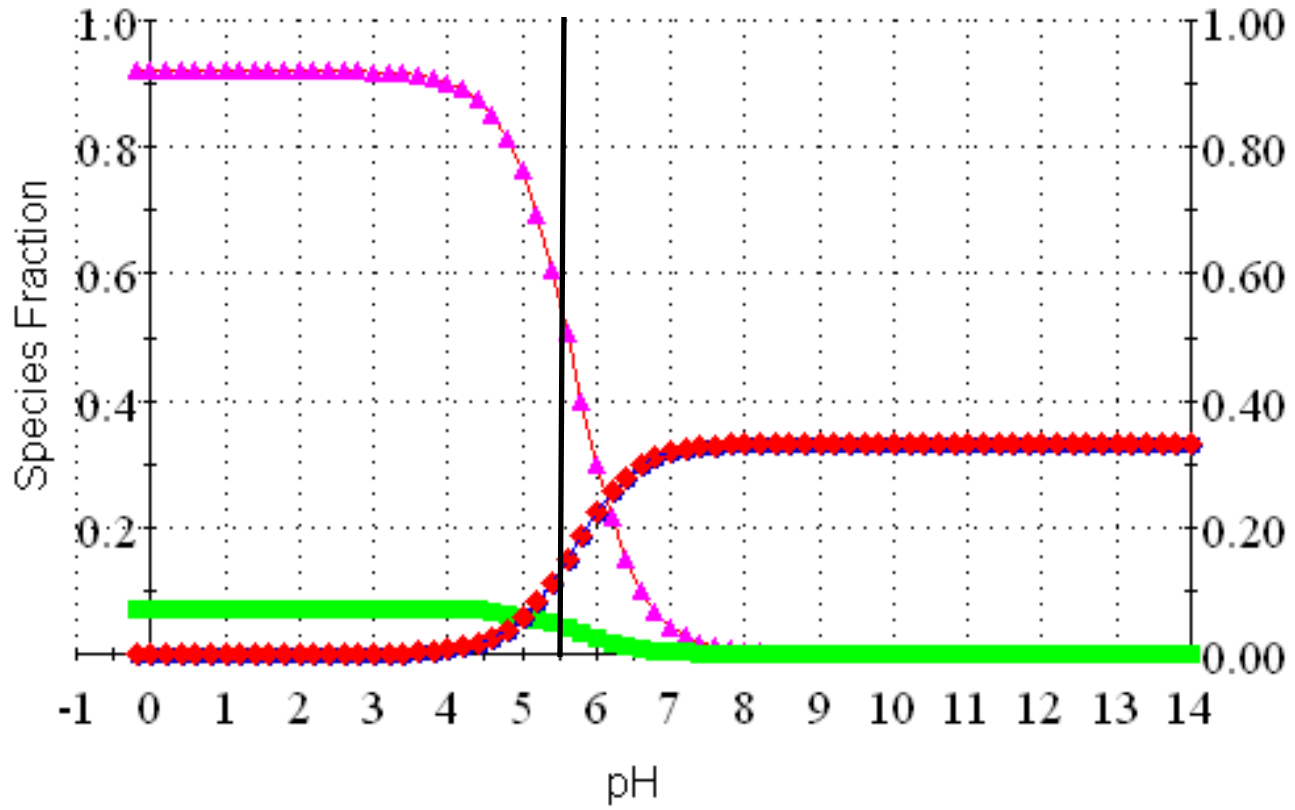


Max Fraction = 0.33

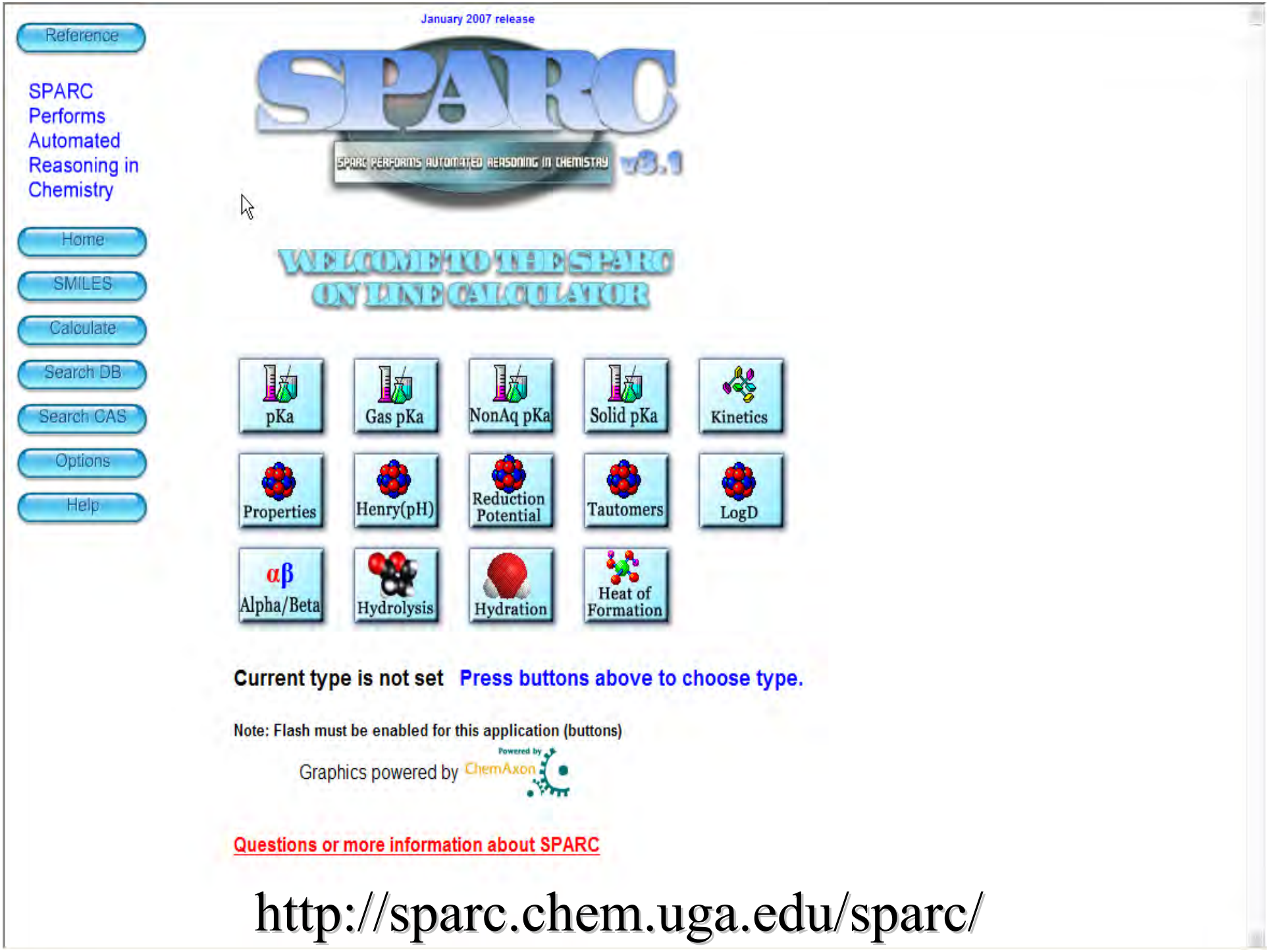


Max Fraction = 0.33

SPARC Speciation Plot

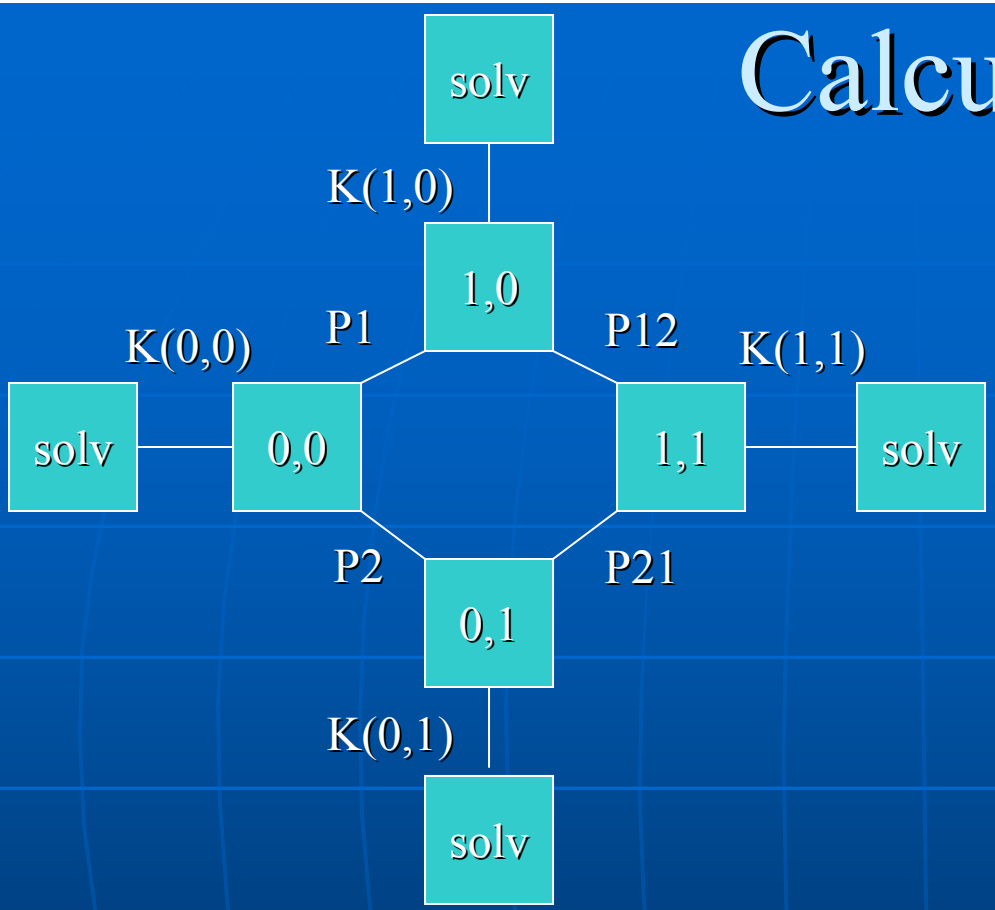


57%		OC(CC(=O)OCC)(O)C(F)(F)F
7%		FC(F)(F)C(=O)CC(=O)OCC
3x12	**36%**	FC(F)(F)C([O-1])=CC(=O)OCC
		FC(F)(F)C(=O)C=C([O-1])OCC
		FC(F)(F)C(=O)[C-1]C(=O)OCC

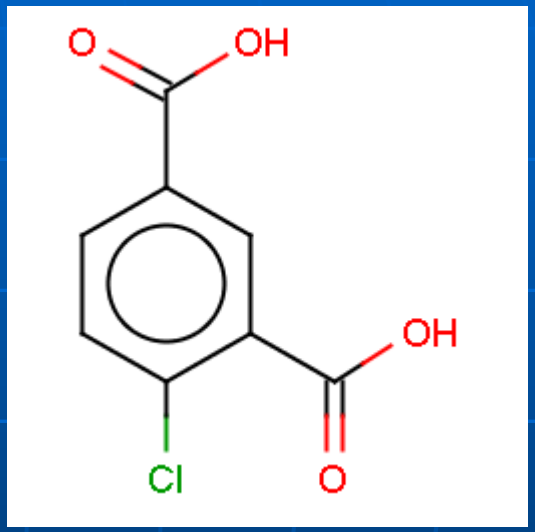


http://sparc.chem.uga.edu/sparc/

Calculating LogD



Let an unionized co2h be 0 and ionized be 1



$\mathbf{D_W = 1 + P1 + P2 + P1 \bullet P12}$ where $P1 = K1[H^+]^{-1}$

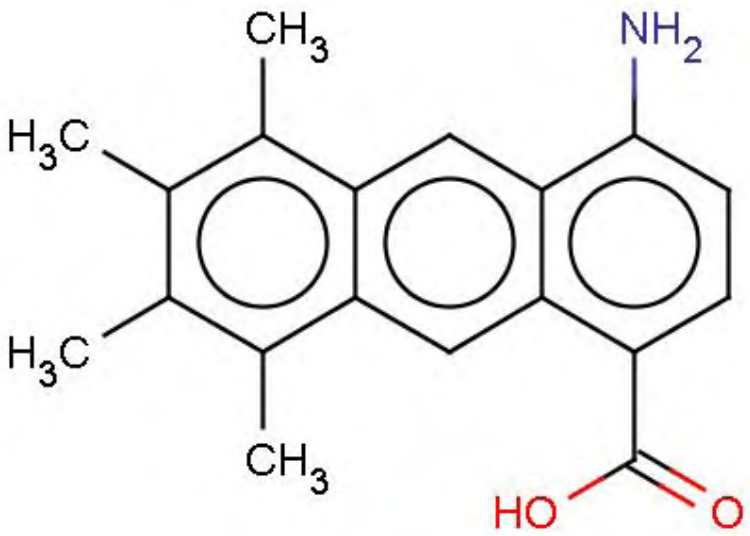
each of the species 0,0, 0,1 etc. are coupled to the octanol phase.

This same distribution in the octanol phase can be represented as

$\mathbf{D_O = K_{OW}(0,0) + K_{OW}(1,0)\,P1 + K_{OW}(0,1)P2 + K_{OW}(1,1)P1 \bullet P12}$

and $\mathbf{LogD = Log(D_O/D_W)}$

LOG D



Log D as a function of pH will be calculated.

Remember to choose a solvent other than water.

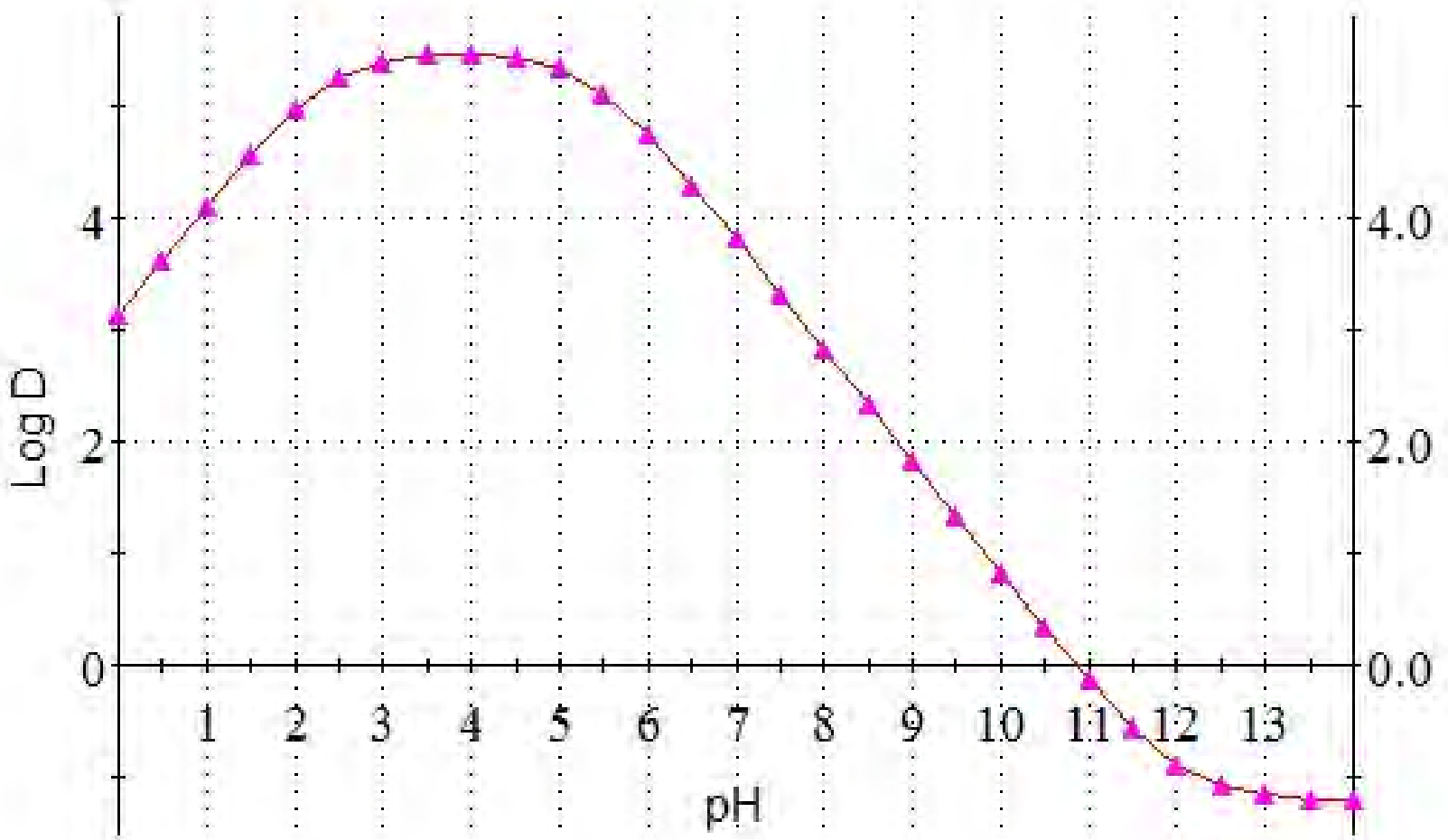
SMILES String
Compound: c1c(N)c2cc3c(C)c(C)c(C)c(C)c3cc2c(C(=O)O)c1

Calculate logD at 25 Degrees C

Calculate logD in the solvent octanol New Solvent

Calculate

SPARC Property Plot

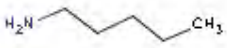


January 2007 release

SPARC

SPARC PERFORMS AUTOMATED REASONING IN CHEMISTRY v3.1

FULL SPECIATION



Options running are:

Total Solid Concentration: 0.001 g/L
Don't consider carbon ionization.

Solid Composition	Solid Properties	Metals
Sand: 100%	Sand Surface Area: 100000 cm^2/g	Charge +1: 0.001 M
Organic: 0%	Organic Type: generic_organic	Charge +2: 0.0 M
Clay: 0%	Clay: Swelling: 0% NonSwelling: 100%	Charge +3: 0.0 M

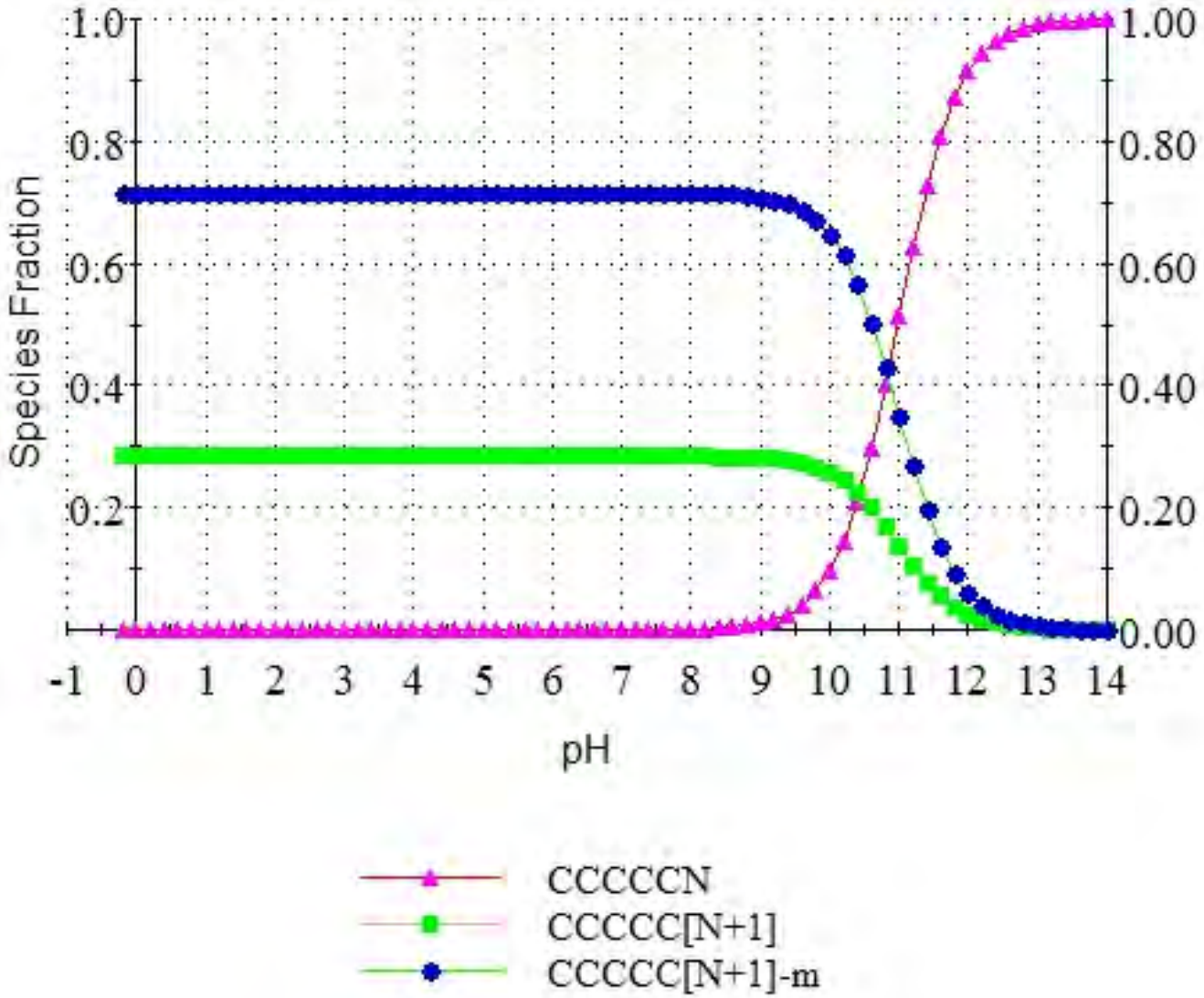
SMILES String:

CCCCCN

This will result in ~1 calculations and may take as long as 0.68 minutes.

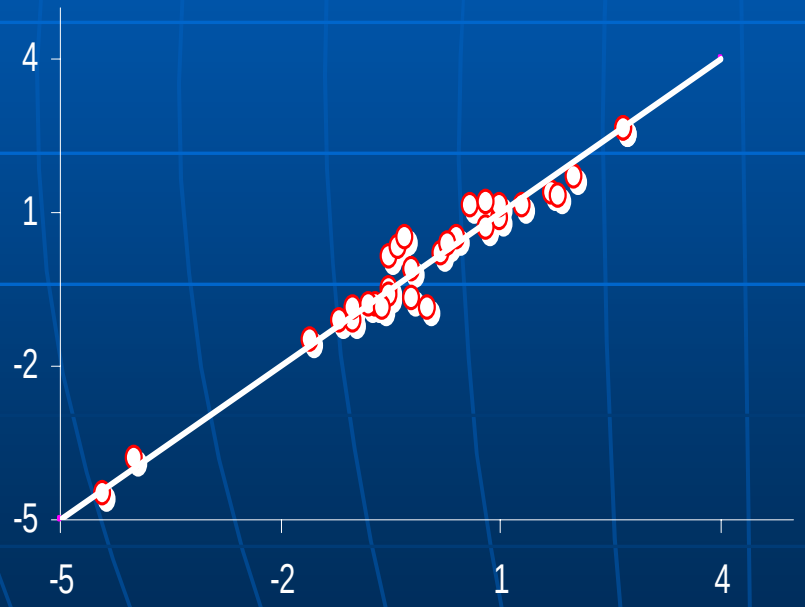
Full Speciation

SPARC Speciation Plot



Hydration and Hydrolysis

31 aldehyde and ketones
RMS 0.34
R^2 0.95



590 Ester Base Hydrolysis
RMS 0.37
R^2 0.94